

EOSC servisi za digitalnu humanistiku

Davor Davidović, Karolj Skala, Peter Kacsuk, Philipp Wieder
davor.davidovic@irb.hr

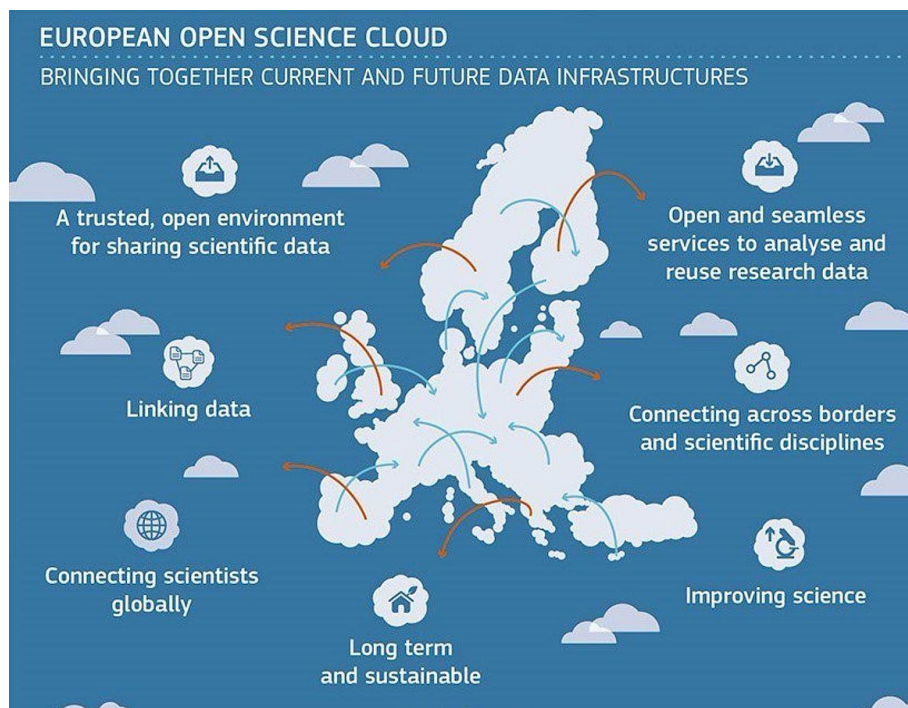


EOSC-hub receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 777536.

Europski Otvoreni Znanstveni Oblak

„The European Open Science Cloud (EOSC) je vizija **federativne**, **globalno dostupne**, **multi-disciplinarne** okoline u kojoj istraživači, inovatori, kompanije i građani mogu međusobno **publicirati**, **pretraživati** i **koristiti** podatke, alate i druge rezultate za svoja istraživanja, inovacije i edukaciju.”

(Open Science Policy Platform EOSC wg)



Izazovi i prepreke

- većina prepreka u realizaciji EOSC ciljeva su više **društvene**, a ne **tehničke** prirode
- glavni tehnički izazov je **kompleksnost podataka i analitičkih procedura** unutar različitih znanstvenih disciplina
- **nedostatak interoperabilnosti** različitih alata, API-a, načina dijeljenja podataka, AAI,...
- **fragmentiranost** – između znanstvenih i ekonomskih domena, nacionalnih i regionalnih granica, upravljačkih modela...
- heterogenost i kratkoročni projektno-orijentirani modeli financiranja (**problem održivosti**)
- zastarjeli sustav financiranja za znanost i inovacije
- podaci dobiveni iz javno financiranih istraživanja nisu uvijek javno dostupni
- nedostatak jasnih prednosti dijeljenja podataka

Digitalna humanistika



digitalizacija
pohrana

Pohrana podataka



Search
Browse
Access
Annotate
Archive

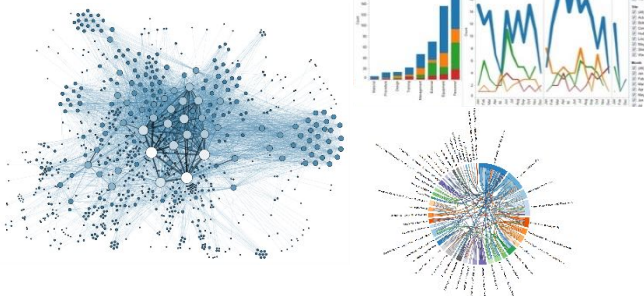
- Podatkovni repozitoriji
- Kolekcije podataka
- Baze podataka
- Arhiviranje
- ...

obrada



Hrvatska

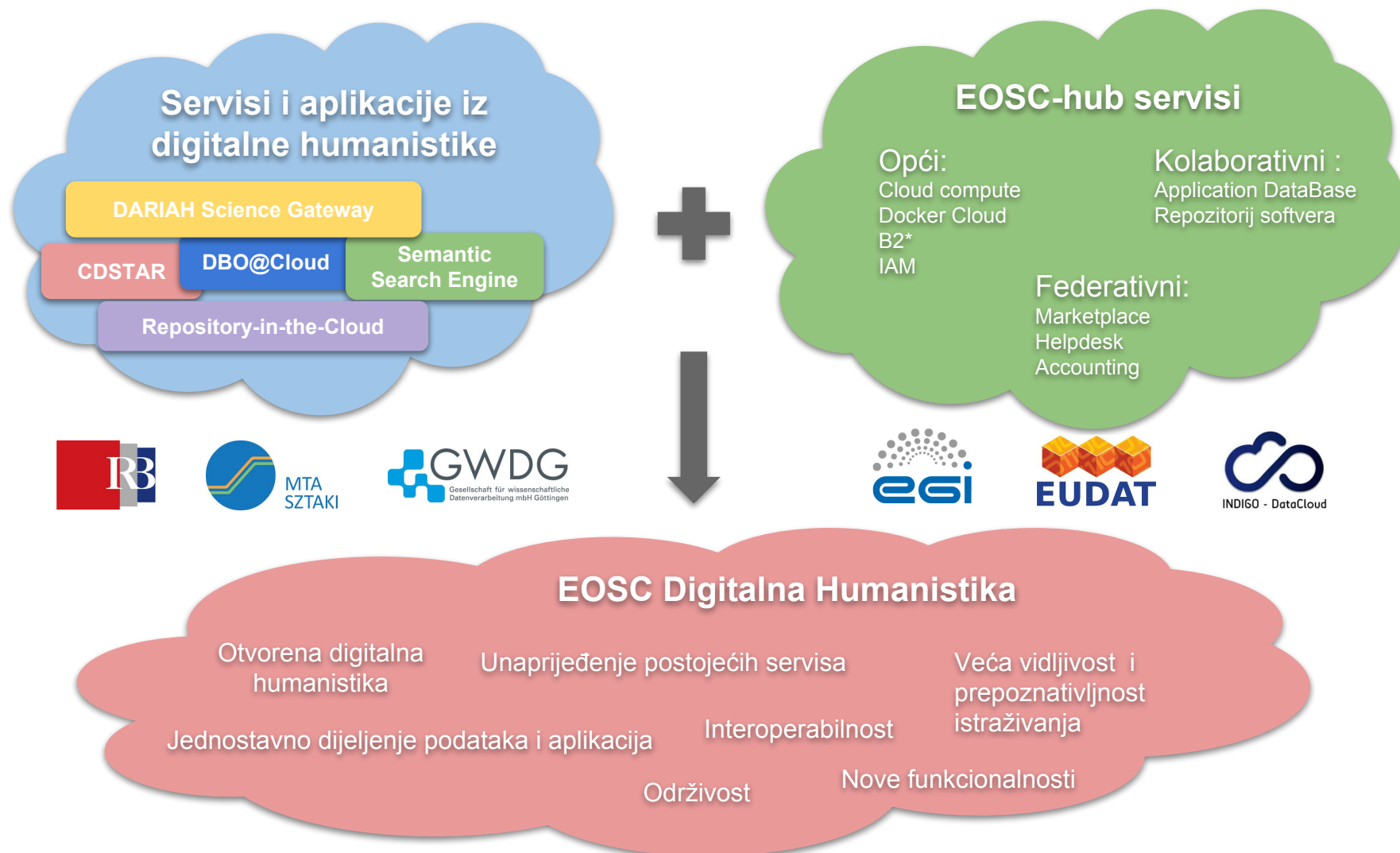
Analiza



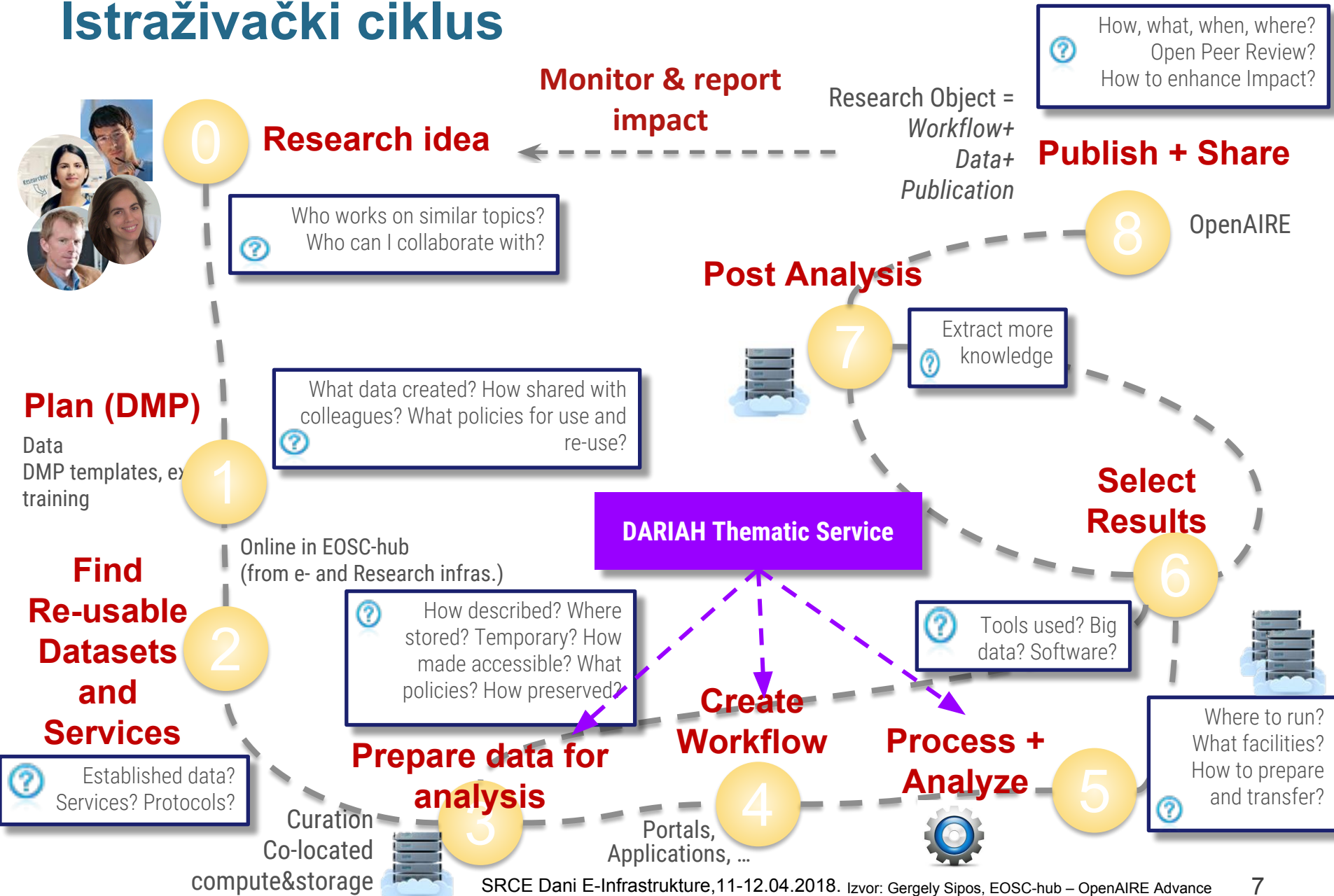
- Antropologija
- Arheologija
- Povijest
- Lingvistika
- Književnost
- Umjetnost
- ...

- Započeo 1.1.2018 (3 godine)
- Cilj je povezati mnogobrojne pružatelje usluga u tzv. Hub
 - centralna točka s koje će istraživači i inovatori moći otkriti, pristupiti i koristiti široki spektar resursa za naprednu podatkovno-orijentiranu znanost
- Objedinjuje ponuđače usluga i servisa: EGI federacija, EUDAT, INDIGO-DataCloud te glavne europske istraživačke infrastrukture (npr. Elixir, CLARIN, DARIAH,...)
- Jedinstveni katalog istraživačkih podataka, servisa i softvera

DARIAH Thematic Service



Istraživački ciklus




Kako EOSC-hub servisi mogu unaprijediti digitalnu humanistiku

Primjeri postojećih servisa

DARIAH znanstveni portal

Link: <https://dariah-gateway.lpds.sztaki.hu/>

← → ↻ Sigurno | <https://dariah-gateway.lpds.sztaki.hu>

 EGI DARIAH CC Gateway Sign In

Welcome Applications Cloud Access Workflow Development File Transfer Help Forum

Welcome

Web Content Display

Welcome to the EGI DARIAH Gateway!

This gateway was developed within the [EGI-Engage H2020 project](#) and operated by the [EGI DARIAH Competence Centre](#). This gateway displays state-of-the-art capabilities of what's available in digital science with the use of EGI. The gateway provides various web-based applications and services for the Digital Humanities researchers, institutes and communities.

The gateway includes a [Help](#) page for giving more information on the applications and services it provides.



The gateway includes the following [applications](#):

- **Simple Semantic Search Engine (SSE):**
Allows users to search in the [e-Infrastructure Knowledge Base](#) (Open Access Document Repositories and Data Repositories). No authentication is required to access the service.
- **Parallel Semantic Search Engine (PSSE):**
PSSE is a parallelised version of SSE enabling simultaneously search across the [e-Infrastructure Knowledge Base](#), [Europeana](#), [Cultura Italia](#), [Isidore](#), [OpenAgris](#), [PubMed](#) and [DBpedia](#) platforms. No authentication is required to access the service.
- **DBO@Cloud:**
A Cloud-based repository presenting 100+ years old collection of Bavarian dialects. No authentication is required to access the service.

The gateway includes the following [services](#):

- **Cloud Access:**
Single-job applications and parameter-sweep applications can be run on the DARIAH VO clouds without porting efforts. [Authentication](#) is required to access the gateway service and then robot certificate provides access to the cloud.
- **Workflow Development:**
Complex workflow applications can be developed and run on all the resources of the DARIAH VO. [Authentication](#) is required to access the gateway service and then robot certificate provides access to the cloud.
- **File transfer:**
Enables transferring data from, to and between storage services providing HTTP, HTTPS, SFTP, GSIFTP, SRM, IRODS and S3 protocols. [Authentication](#) is required to access the gateway service and then additional authentication is needed to access the selected storage service.

Web Content Display

  DARIAH Science Gateway is an EGI service provided by DARIAH Competence Centre, co-funded by EGI-Engage <http://go.egi.eu/eng>

DARIAH znanstveni portal

Prednosti korištenja EOSC servisa:

- Usluga smještaja u računalnom oblaku - povećanje dostupnosti i skalabilnosti (EGI Cloud service)
- Jedinstvena prijava korištenjem npr. EduHr (AAI, IAM)
- Praćenje prometa i statistike korištenja portala (EGI Accounting)
- Jednostavno pokretanje vlastite instance portala unutar oblaka (App Database, Marketplace)
- Tehnička korisnička podrška (Helpdesk platform)

Podatkovni repozitorij u oblaku

- Rezultat Indigo-DataCloud projekta
- Temeljen na Invenio platformi
- Omogućuje brzo stvaranje vlastitog podatkovnog repozitorija u računalnom oblaku
- Nove funkcionalnosti integracijom EOSC servisa
 - Pristup putem postojećih korisničkih računa, npr. EduHr (AAI)
 - Korištenje računalne i podatkovne infrastrukture u oblaku (EGI Cloud Service)
 - Jedinstveni opis traženih računalnih resursa i rezervacija infrastrukture (TOSCA Orchestrator)
 - Automatska izrada i pohrana sigurnosne kopije čitavog repozitorija (B2SAFE)



Semantička pretraga baze podataka

The image shows a web browser window displaying the EGI DARIAH CC Gateway. The search results for 'bellis perennis' are shown, with a list of records found. The first record is highlighted, showing the title 'Prvo priopćenje o vrsti Puccinia distincta McAlpine, novoj europskoj hrđi na tratinčicama (Bellis perennis L.) iz Hrvatske (Citations) (Linked Data)' and the author 'Weber Roland W.S., Lehrbereich Biotechnologie Universität Kaiserslautern Paul-Ehrlich-Str.23 67663 Kaiserslautern Germany; Jurc Dušan; Gozdarski Institut Slovenije Vehnapot2 1000 Ljubljana p.p.2985 Slovenia'.

The detailed view of the article from the journal '10hrčak' is shown. The article title is 'Prvo priopćenje o vrsti Puccinia distincta McAlpine, novoj europskoj hrđi na tratinčicama (Bellis perennis L.) iz Hrvatske'. The authors are 'Dušan Jurc' and 'Roland W.S. Weber'. The article is published in 'Natura Croatica, Vol.9 No.4 Prosinac 2000.'.

The article abstract (Sažetak) states: 'Teška infekcija hrđe na divljim i uzgajanim tratinčicama (Bellis perennis L.) otkrivena je u travnju i svibnju 2000. na izoliranom probalnom području između Lovrana i Opatije. Patogena gljiva bila je determinirana kao Puccinia distincta McAlpine, koja se protekle četiri godine brzo širi po Europi, ...'. The article is available in PDF format (479 KB) and can be downloaded from the journal's website.

The article is also available in English (Engleski) and can be downloaded from the journal's website. The article is available in PDF format (479 KB) and can be downloaded from the journal's website.

Analiza e-knjiga iz kolekcije Gutenberg

- Jednostavna analiza teksta svih e-knjige određenog autora objavljenih u kolekciji Gutenberg - <https://www.gutenberg.org>
- Dijelovi analize (workflow):
 - Spojiti se na bazu Gutenberg
 - Za danog autora (ime i prezime) skinuti sve objavljene e-knjige
 - “Očistiti” tekst od suvišnih dijelova (npr. zaglavlja)
 - Analizirati skinute tekstove (python programski jezik)
 - Frekvencija riječi
 - Broj tokena
 - Lingvistička raznolikost
 - Naći riječi duže do 7 znakova koje se pojavljuju barem 10 puta
- Kako pokrenuti analizu:
<https://bitbucket.org/davordavidovic/textanalysis/overview>

Pretraga DBÖ kolekcije (DBÖ@Cloud)

- 100 godina stara baza podataka Bavarskih dijalekata iz doba Austro-ugarske monarhije (podaci od Austrijske akademije znanosti)
- 50,000+ zapisa
- Baza je pohranjena u računalnom oblaku na više lokacija

DBÖ@Cloud

Browser address bar: <https://dariah-gateway.lpd.sztaki.hu/web/guest/dbo-portlet>

Navigation bar: Admin | Davor Davidovic

EGI DARIAH CC Gateway

Applications / DBO Portlet

DARIAH Browser

Titel	Autor
Die Moosburg Arnulfs von Kärnten	Svetina, Robert
@cs Das Jagdbuch des Martin Strasser von Kollnitz	
Usi linguistici al femminile nella ricerca linguistica a Sappada (per il dizionario sappadino)	Hornung, Maria
Das besondere Museum	Hornung, Maria
Ladinia et Romania	
Personennamen und -identität	
Ir sult sprechen willekommen	
Lexikalische Studien zum Dialekt im Wiener Volksstück vor Nestroy	Reutner, Richard
Die Mundartdichtung in Salzburg	Rettenbacher, August/Rettenbacher, Barbara
Die Überlieferung der Ranshofener Traditionen	Schmidt, Rudolf Wolfgang
Beiträge zur älteren Besitz- und Rechtsgeschichte steirischer Klöster	Pirchegger, Hans
Geschichte des Bezirks Gröbmring	Pirchegger, Hans
Burgengründer	Neweklowsky, Walter
Die Linzer Fernstrassen	Pfeffer, Franz
Dizionario toponomastico del Friuli-Venezia Giulia	Frau, Giovanni
Die Kompositen in gallischen Personennamen	Schmidt, Karl Horst
Gaulish Personal Names	Evans, D. Ellis
Das ehemalige Cisterciensinnen-Kloster zu St. Bernhard bei Horn	Endl, Friedrich
Salzburg und Hamburg Erzbistumsgründung und Missionspolitik in karolingischer Zeit	Wavra, Brigitte
Karantanien und der Alpen-Adria-Raum im Frühmittelalter	
Quellenverzeichnis zu den mittelhochdeutschen Wörterbüchern	Nellmann, Eberhard
---	Anreiter, Peter
Necrologium Curiense	
Die Gründung des Klosters Maria-Zell im Wiener Wald und die Besitzgeschichte seiner Stifterfamilie	Lechner, Karl
Schriftsprache und Dialekte im Deutschen nach Zeugnissen alter und neuer Zeit	Socin, Adolf
Althochdeutsches Glossenwörterbuch	Wells, John C.
Régi Magyar Családnevek Szótára	Miklós, Kázmér

Category	File Type	File Name	Container Path	File Size	Uploa...
----------	-----------	-----------	----------------	-----------	----------

Id:	32
Titel:	Schriftsprache und Dialekte im Deutschen nach Zeugnissen alter und neuer Zeit
Autor:	Socin, Adolf
Untertitel:	Beiträge zur Geschichte der deutschen Sprache
Nachdruck:	Reprografischer Nachdruck der Ausgabe Heilbronn 1888
Erscheinungsort:	Hildesheim, New York
Erscheinungsjahr:	1970
Erstauflage:	1888
Verlag:	Georg Olms Verlag
Seitenstruktur:	XII+544
Kurzzitat Sort:	-

DBÖ@Cloud

Applications / DBO Portlet

DARIAH Browser

Sign In

Titel	Autor	Kurzzitat	Langzitat
Die Woschung Altmanns von Kärnten	Svetina, Robert		
@cs Das Jagdbuch des Martin Strasser von Kollnitz		Strasser (1624/1997)	Das Jagdbuch des Martin Strasser von Kollnitz
Lex. linguistici si femminile nella ricerca linguistica a Sappada (per il dizionario eppo)	Hornung, Maria		
Das besondere Museum	Hornung, Maria		
Ladinia et Romania			
Personennamen und -identität			
Ir sult sprechen willekommen			
Lexikalische Studien zum Dialekt im Wiener Volksstück vor Nestoy	Reutner, Richard		
Die Mundartdichtung in Salzburg	Rettenbacher, August/Rettenbacher, Barbara		
Die Überlieferung der Ranshofener Traditionen	Schmidt, Rudolf Wolfgang		
Beiträge zur älteren Besitz- und Rechtsgeschichte steirischer Klöster	Pirchegger, Hans		
Geschichte des Bezirks Gröbmung	Pirchegger, Hans		
Burgengründer	Neweklowsky, Walter		
Die Linzer Fernstrassen	Pfeffer, Franz		
Dizionario toponomastico del Friuli-Venezia Giulia	Frau, Giovanni		
Die Kompositen in gallischen Personennamen	Schmidt, Karl Horst		
Gaulish Personal Names	Evans, D. Ellis		
Das ehemalige Cisterciensinnen-Kloster zu St. Bernhard bei Horn	Endl, Friedrich		
Salzburg und Hamburg Erzbistumsgründung und Missionspolitik in karolingischer Zeit	Wavra, Brigitte		
Karantien und der Alpen-Adria-Raum im Frühmittelalter			
Quellenverzeichnis zu den mittelhochdeutschen Wörterbüchern	Nellmann, Eberhard		
---	Anreiter, Peter		
Necrologium Curiese			
Die Gründung des Klosters Maria-Zell im Wiener Wald und die Besitzgeschichte sein...	Lechner, Karl		
Schriftsprache und Dialekte im Deutschen nach Zeugnissen alter und neuer Zeit	Socin, Adolf		
Althochdeutsches Glossenwörterbuch	Wells, John C.		
Régi Magyar Csaldnevek Szótára	Miklós, Kázmér		
Krankheit und Heilen	Grabner, Elfriede		
Zur Textgeschichte der Vita Ruperti	Beumann, Helmut		
Wie alt ist Klagenfurt?	Webering, Evelyn		
Namenkundliches aus dem Bereich der Landeshauptstadt Klagenfurt	Pohl, Heinz Dieter		
Lexikon deutschsprachiger Mundartautoren	Sowinski, Bernhard	SOWINSKI: (1997)	
Dizionario di Toponomastica	Gasca Queirazza, Giuliano/Marcatto, Carla/Pellegrini, Giov...		
Österreichische Redewendungen und Redensarten	Malygin, Viktor T.		

Detalji zapisa

Id:	2
Typ id:	1
Veroeffentlicht:	1
Titel:	@cs Das Jagdbuch des Martin Strasser von Kollnitz
Band:	3. Bd.
Reihentitel:	Das Kärntner Landesarchiv. Studien zur geschichtlichen Landeskunde Kärntens
Auflage:	2., unveränd. Aufl.
Erscheinungsort:	Klagenfurt
Erscheinungsjahr:	1997
Verlag:	Verlag des Kärntner Landesarchivs
Isbn:	3-900531-35-8
Seitenstruktur:	419
Herausgeber:	Lindner, Kurt
Kurzzitat:	Strasser (1624/1997)
Kurzzitat Sort:	Strasser (1624/1997)
Langzitat:	Das Jagdbuch des Martin Strasser von Kollnitz. Hg. von Lindner, Kurt. 2., unveränd. Aufl. Klagenfurt 1997 (= Das Kärntner Landesarchiv. Studien zur geschichtlichen Landeskunde Kärntens 3)
Freigabe:	1
Checked:	1
Online:	1
Anmerkung:	ANM_BIZ: aus dem Jahre 1624 (Ain Puech von allerlei Jägerei und Waldmanschaften). ANM_EXZ: Exz 0 == 419 Seiten == Lf: Gewerkensohn aus Gastein Sa., oberster sa. Jägermeister, 1600 nach Kollnitz in Gm. St. Paul Kä. auf Erzbisch. Besitz ANM_S: KÜ.2002/01: Eintr.ok ## Sig. ergänzen. Geb.Zuordng.??, literarisch? M.G.: SL/A: 2./4. Anmerkung: weitere Bände des "Kärntner Landesarchivs" unter N-8189 und A-7135

Category	File Type	File Name	Container Path	File Size	Upload...
No items to show.					

Zaključak

Europska otvorena znanost

- dijeljenje i ponovno korištenje znanstvenih podataka, računalnih servisa i aplikacija
- interdisciplinarnost
- povezivanje svih dionika znanstvenog procesa (istraživači, ponuđači servisa i podatkovnih infrastruktura...)

Digitalna humanistika

- **otvaranje** znanstvenih podataka (učiniti podatke dostupnijima široj zajednici) i interoperabilnost
- **razvoj** novih servisa i softvera temeljenih na naprednim (oblačnim) računalnim tehnologijama



Hvala!

Davor Davidović

davor.davidovic@irb.hr

EOSC-Hub DARIAH Thematic Service &
DARIAH Competence Centre



eosc-hub.eu



@EOSC_eu



eosc-hub-project